

## ON THE LUMINOSITY DEPENDENCE OF THE GALAXY PAIRWISE VELOCITY DISPERSION

JEREMY L. TINKER<sup>1</sup>, PEDER NORBERG<sup>2</sup>, DAVID H. WEINBERG<sup>3</sup> & MICHAEL S. WARREN<sup>4</sup>*Draft version February 4, 2008*

## ABSTRACT

We make predictions for the pairwise velocity dispersion (PVD) of galaxies with models that are constrained to match the projected correlation function and luminosity function of galaxies in the Two-Degree Field Galaxy Redshift Survey (2dFGRS). We use these data to constrain the halo occupation distribution (HOD) of 2dFGRS galaxies, then calculate the PVD by populating the halos of a large-volume, high-resolution N-body simulation. We examine the luminosity and scale dependence of the predicted PVD. At large scales the PVD is flat, but at  $r \sim 2 h^{-1}$  Mpc the dispersion rapidly rises as the dominant source of galaxy pairs transitions from two-halo pairs to pairs within a single halo. At small and large scales,  $r < 1 h^{-1}$  Mpc and  $r \gtrsim 3 h^{-1}$  Mpc, we find that the PVD decreases with increasing galaxy luminosity. This result is mostly driven by the fraction of satellite galaxies  $f_{\text{sat}}$ , which is well-constrained by the shape and amplitude of the correlation function. We find  $f_{\text{sat}} \sim 25\%$  for galaxies fainter than  $L_*$ , while for brighter galaxies the satellite fraction rapidly declines, creating the decrease in the PVD with luminosity. At  $r = 1 h^{-1}$  Mpc, the PVD has no dependence on luminosity because satellite galaxies dominate the statistics for both bright and faint objects. Recent measurements of the PVD in Fourier space using the “dispersion model” have reported a strong decline in PVD with increasing luminosity at  $k = 1 \text{ Mpc}^{-1} h$ . We test this method with our HOD models, finding that the dispersion model roughly recovers the shape of PVD with scale, but that there is no consistent comparison between the PVD at a given  $k$ -scale and the true dispersion at a given value of  $r$ . This results in a luminosity dependence in  $k$ -space that is stronger than in configuration space. The luminosity dependence of the HOD results in Fourier space are consistent with those measured at  $k = 1 \text{ Mpc}^{-1} h$ ; thus the recent measurements of the PVD are fully explainable in the context of halo occupation models. The normalization of the PVD is lower than predicted by our fiducial model, and reproducing it requires a lower value of  $\Omega_m$  ( $\sim 0.2$  instead of 0.3), a lower value of  $\sigma_8$  ( $\sim 0.7$  instead of 0.9), or a significant bias between the velocity dispersion of galaxies and dark matter in massive halos.

*Subject headings:* cosmology:theory — galaxies:halos — large scale structure of the universe

## 1. INTRODUCTION

Galaxy peculiar velocities, their motions with respect to the smooth Hubble flow, offer a unique window into both cosmology and galaxy formation. The galaxy velocity field reflects the matter density of the universe,  $\Omega_m$ , and the amplitude of matter fluctuations (Peebles 1976; Sargent & Turner 1977; Kaiser 1987). This is true at large scales, where the infall velocities of galaxies toward overdense regions are determined by the overall amount of matter and its distribution at large scales, and it is also true at small scales, where the random motions of galaxies are determined by the shape and normalization of the mass function of dark matter halos (Sheth 1996). But peculiar velocities also provide information about galaxy formation and evolution; just as galaxies are a biased tracer of the spatial distribution of dark matter, they can also be biased with respect to the velocity distribution of dark matter (Berlind et al. 2003; Yoshikawa et al. 2003; Faltenbacher et al. 2005). Pairwise velocity statistics generally differ even if galaxies have the same peculiar velocities as local dark matter particles because spatial bias changes the weighting of those velocities. This is of particular importance at small scales, reflecting the accretion histories and evolution of galaxies within groups and clusters. A full model of galaxy formation must reproduce both the real-space clustering of the observed galaxy distribution and the distribution of velocities, and the dependence of these distributions on lu-

minosity and galaxy type.

In this paper we make predictions for the pairwise velocity dispersion of galaxies (hereafter PVD, or  $\sigma_v$ ) based on the observed real-space clustering of galaxies in the Two-Degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001). We use the Halo Occupation Distribution (HOD) modeling method, in which the relation between galaxies and mass is described at the level of virialized dark matter halos (Kauffmann et al. 1997; Jing et al. 1998; Benson et al. 2000; Seljak 2000; Peacock & Smith 2000; Ma & Fry 2000; Scoccimarro et al. 2001; Berlind & Weinberg 2002). The mass function and spatial clustering of dark matter halos themselves, for given cosmological parameters, are well-determined from numerous numerical and analytic studies (e.g., Sheth & Tormen 1999; Jenkins et al. 2001; Sheth et al. 2001; Seljak & Warren 2004; Tinker et al. 2005b; Warren et al. 2005). In our approach, the parameters that determine the relationship between halo mass and the average number of galaxies within a halo are constrained by observational measurements of the projected two-point correlation function  $w_p(r_p)$  (see, e.g., Zheng 2004; Zehavi et al. 2004, 2005; Tinker et al. 2005b). This technique is similar to the conditional luminosity function (CLF) analysis of Yang et al. (2003), but utilizes spatial clustering over a wide range of scales, linear to non-linear. Once the parameters of the occupation function have been determined, predictions can be made for other clustering statistics either analytically or through the creation of mock galaxy distributions with numerical simulations. In this paper we obtain the PVD by populating the halos of a high-resolution collisionless N-body simulation with galaxies such that both the spatial clustering

<sup>1</sup> Kavli Institute for Cosmological Physics, University of Chicago

<sup>2</sup> Institute for Astronomy, University of Edinburgh

<sup>3</sup> Department of Astronomy, Ohio State University

<sup>4</sup> Theoretical Astrophysics, Los Alamos National Labs

as a function of luminosity and the galaxy luminosity function match observations. This approach complements the analytic model for the PVD of Slosar et al. (2006), and our qualitative conclusions about the PVD are consistent with their analysis.

Directly measuring peculiar velocities requires knowledge of the true distance to galaxies, thereby removing the Hubble flow from the measured redshift. This limitation subjects these measurements to shot noise and sample variance. To circumvent this problem, many studies have used measurements of the two-point galaxy correlation function in redshift space  $\xi(r_p, r_\pi)$ , which is distorted along the line of sight by the peculiar velocities (e.g., Davis & Peebles 1983; Bean et al. 1983). In this method, the galaxy PVD is obtained by modeling the redshift-space correlation function at a given projected separation  $r_p$ , assuming that the PVD is a constant along the line of sight. This method has been used in more recent measurements of the galaxy PVD by Jing et al. (1998), Zehavi et al. (2002), and Hawkins et al. (2003). We will compare our predictions to the 2dFGRS results of Hawkins et al. (2003), demonstrating qualitative agreement between the measurements and a model normalized to a matter density of  $\Omega_m = 0.2$ . A quantitative comparison is difficult due to the flux-limited sample used in Hawkins et al. (2003), and to possible systematic errors of the model used to infer the PVD from the observed  $\xi(r_p, r_\pi)$ .

Recently, Jing & Börner (2004, hereafter JB) and Li et al. (2005, hereafter L05) have presented measurements of the PVD obtained by modeling observations of the redshift-space correlation function for volume-limited samples in the 2dFGRS (JB) and Sloan Digital Sky Survey (L05). One notable result of these two studies is the luminosity dependence of the PVD at small scales. Both groups measure a clear trend of decreasing velocity dispersion with increasing luminosity, until approximately  $M_* = -1$ , whereupon the PVD rapidly rises. JB claim that this luminosity dependence is counterintuitive, and both papers show it is not reproduced in a halo occupation model.

The technique of JB and L05 is to infer the PVD from redshift-space clustering in Fourier space using the “dispersion model” of the redshift-space power spectrum (e.g., Cole et al. 1994, 1995; Peacock & Dodds 1994; Hatton & Cole 1998, 1999; see Scoccimarro 2004 for a thorough review and critique of this model). In their implementation of the dispersion model, JB and L05 use the redshift-space power spectrum at a fixed  $k$ -scale and fit the data as a function of angle, as opposed to a constant projected separation. The meaning of the velocity dispersion obtained in this case is less clear. JB argue that the qualitative transformation between the PVD in Fourier space and that in configuration space should be  $k \leftrightarrow 1/r$ , an inference motivated by the transition between one-halo and two-halo clustering, which appears to occur at  $1 \text{ Mpc}^{-1} h$  in  $k$ -space and  $1 h^{-1} \text{ Mpc}$  in configuration space. The  $k \leftrightarrow 1/r$  transformation is also supported by tests with mock galaxy distributions (Jing & Börner 2001). JB and L05 compare their results to an implementation of a CLF fit to the 2dFGRS of Yang et al. (2003). This CLF model predicts a PVD that increases with luminosity, nearly orthogonal to the observed trend.

In this paper we test the JB implementation of the dispersion model against our HOD models. We find that the Fourier space dispersion measured by JB on  $k = 1 \text{ Mpc}^{-1} h$  can, in its interpretation as a configuration space dispersion on  $r = 1 h^{-1} \text{ Mpc}$ , be systematically off the true configuration space value at the 20–40% level, and only roughly follows the true

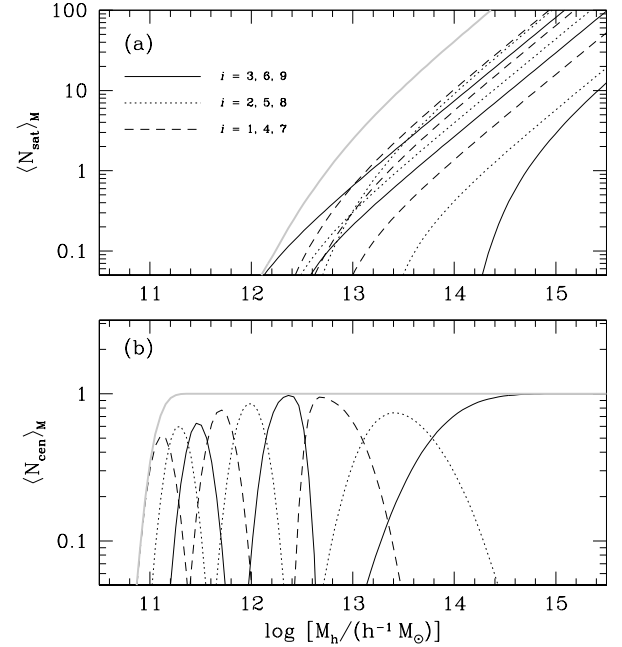


FIG. 1.— Panel (a): Satellite occupation functions for all nine magnitude bins, as listed in Table 1. The curves go in order of faintest to brightest galaxies from left to right. The thick gray line shows the total satellite occupation function for all galaxies. Panel (b): Central occupation functions for the same magnitude bins. The ordering of the lines is the same as in panel (a). The thick gray line is the sum of all central occupation functions.

dependence of the PVD with configuration space scale. We pay particular attention to the question of what physical scale the  $k = 1 \text{ Mpc}^{-1} h$  measurements are probing. In configuration space, our HOD models predict that  $\sigma_v$  should be nearly independent of luminosity at  $r = 1 h^{-1} \text{ Mpc}$ . At smaller and larger scales,  $\sigma_v$  has a weak dependence on luminosity, such that brighter galaxies have lower dispersions. Applying the dispersion model technique to our mock galaxy distributions, we demonstrate that the  $k = 1/r$  comparison is reasonable for only a small set of luminosity samples, and in general there is no consistent comparison between dispersion model results and the true PVD at a given scale. Thus in Fourier space the dependence of the PVD on galaxy luminosity is steeper than that in configuration space, leading to the strong trend measured in the JB and L05 results.

## 2. DATA AND MODELING

### 2.1. 2dFGRS Data

We make predictions for the PVD and its dependence on luminosity by inferring the mean halo occupation function from measurements of the projected galaxy correlation function  $w_p(r_p)$  of the 2dFGRS. These measurements are similar to those presented in Norberg et al. (2001) and Norberg et al. (2002a), who measured  $w_p(r_p)$  from a sample of  $\sim 160,000$  redshifts, but our data have been updated to include the full data release of the 2dFGRS, which contains  $\sim 221,000$  redshifts. Further details about the survey can be found in Colless et al. (2001, 2003) and Norberg et al. (2002b).

The details of the clustering measurements can be found in Norberg et al. (in preparation). We present here a brief summary of the calculations. Using the full 2dFGRS survey we create volume limited samples, with faint limits from  $M_{b_j} =$

$-17.0$  to  $M_{b_j} = -21.0$ , each sample  $0.5$  magnitudes<sup>5</sup> wide, for a total of nine samples. All galaxies brighter than  $M_{b_j} = -21$  are grouped into a single sample. As in Norberg et al. (2001, 2002a), a careful account of the selection function is made and the correlation functions are obtained using the standard Landy & Szalay (1993) and Hamilton (1993) estimators, with typically 100 times more randoms than galaxies. The projected correlation function is estimated by integrating  $\xi(r_p, r_\pi)$  out to  $r_{\pi, \max} = 70 h^{-1} \text{Mpc}$ , providing a stable estimate for  $w_p(r_p)$  out to at least  $r_p = 40 h^{-1} \text{Mpc}$ . Due to the sensitivity of the results on close pair incompleteness, clustering results from 2dFGRS are only really reliable beyond typically  $\sim 150 h^{-1} \text{kpc}$ , which is the limit adopted in this paper too. The correlation function is measured in thirteen radial bins, spaced evenly by  $0.2$  in  $\log_{10} r$  beginning at  $\log_{10} r = -0.7$ .

The errors on the clustering measurements are estimated by a bootstrap resampling technique on roughly equal sized subregions, of which there are 16 in total (8 in each 2dFGRS region, covering in each region approximately the same survey area). The main reason for choosing this technique over mock surveys is that it allows error estimates to be obtained for samples with different clustering properties, whereas current available mocks need to be tuned for each sample. Moreover the existence of large structures in the survey (see, e.g., Baugh et al. 2004) are difficult to account for in the available mocks, but with bootstrap resampling of large subregions this type of cosmic variance can, to some extent, be accounted for. We create 100 bootstrap resamplings, which is sufficient to obtain a stable error covariance matrix for each subsample (see Porciani & Norberg 2006 for details on stability tests for error matrices from bootstrap resamplings). We note though that this number of bootstrap resamplings is not large enough to address the covariance between the subsamples. For that reason, in this paper, we have to ignore the existing covariance between ‘neighboring’ volume limited samples. We use principal component analysis of the covariance matrix to calculate the  $\chi^2$  of a given HOD model (e.g., Porciani & Giallisco 2002), using the first twelve principal components.

## 2.2. HOD Models

To model  $w_p(r_p)$ , we use the analytic model described in detail in Tinker et al. (2005b) (see also, Berlind & Weinberg 2002; Cooray & Sheth 2002; Zheng 2004). The mean occupation function is divided into two components; galaxies located at the center of mass of the host halo,  $\langle N_{\text{cen}} \rangle_M$ , and satellite galaxies distributed throughout the host halo with (usually) the same distribution as the dark matter,  $\langle N_{\text{sat}} \rangle_M$ . For galaxy samples defined by a luminosity or magnitude threshold,  $\langle N_{\text{cen}} \rangle_M$  can be defined by a single halo mass scale below which halos are not able to contain galaxies in the sample. For magnitude bin samples,  $\langle N_{\text{cen}} \rangle_M$  must have both a minimum and a maximum mass scale, the latter being defined by the minimum mass of the next brightest sample. A simple example would be to use a square window as the central occupation function, such that  $\langle N_{\text{cen}} \rangle_M^i = 1$  for  $M_{\min}^i < M_h < M_{\min}^{i+1}$ , where  $i$  denotes magnitude bin. This parameterization assumes a unique mapping of central galaxy luminosity onto halo mass. A more physical model takes into account the scatter between mass and luminosity, such as that used by Zheng et al. (2005) to model the occupation functions of semi-analytic and hydro-

dynamic simulations of galaxy formation. Here we modify their  $\langle N_{\text{cen}} \rangle_M$  to take into account binned magnitude samples, i.e.

$$\begin{aligned} \langle N_{\text{cen}} \rangle_M^i &= \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\min}^i}{\sigma_{\log M}^i} \right) \right] - \langle N_{\text{cen}} \rangle_M^{i+1}, \quad 1 \leq i \leq 8, \\ \langle N_{\text{cen}} \rangle_M^i &= \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\log M - \log M_{\min}^i}{\sigma_{\log M}^i} \right) \right], \quad i = 9, \end{aligned} \quad (1)$$

where  $\langle N_{\text{cen}} \rangle_M^i$  represents central galaxies in magnitude bin  $i$ ,  $M_{\min}^i$  is the cutoff mass scale for central galaxies,  $\sigma_{\log M}^i$  controls the width of the cutoff mass range, and  $\text{erf}$  is the error function. For luminosity threshold samples,  $M_{\min}$  is defined as the mass at which  $\langle N_{\text{cen}} \rangle_M = 0.5$ , but in equation (1) this mass can differ from  $M_{\min}^i$ . The form of equation (1) guarantees that the sum of  $\langle N_{\text{cen}} \rangle_M^i$  over all  $i$  will never be larger than one, which would be unphysical. The form we use for the satellite galaxy occupation function is

$$\langle N_{\text{sat}} \rangle_M^i = \exp \left( -\frac{M_{\text{cut}}^i}{M - M_{\min}^i} \right) \left( \frac{M}{M_1^i} \right)^{\alpha_{\text{sat}}}, \quad (2)$$

where  $M_{\text{cut}}^i$  is the cutoff mass scale at which  $\langle N_{\text{sat}} \rangle_M^i$  transitions between a linear dependence of  $N$  on  $M$  and an exponential cutoff,  $M_1^i$  is the amplitude of the occupation function (for  $M_1^i \gg M_{\text{cut}}^i$ ,  $M_1^i$  is the mass at which halos host on average one satellite of magnitude  $i$ ), and  $\alpha_{\text{sat}}$  is the power-law dependence of  $\langle N_{\text{sat}} \rangle_M^i$  on host mass. We assume the scatter about the mean follows a Poisson distribution (Zheng et al. 2005; Kravtsov et al. 2004). Although  $\alpha_{\text{sat}}$  is a free parameter, we assume that it is constant for all magnitude bins. Allowing  $M_{\text{cut}}^i$  to vary as a function of magnitude allows enough flexibility in the satellite occupation function to model  $w_p(r_p)$  for each bin.

For spatial clustering, differentiating between central and satellite galaxies is important because pairs of galaxies that involve a central and a satellite galaxy have a different distribution in the small-scale (i.e. ‘one-halo’) regime than satellite-satellite pairs. This differentiation is also important when modeling galaxy velocities. Central galaxies are assumed to be at rest with the center of mass of the halo, while satellites have random motions roughly equal to the virial dispersion of the dark matter halo. Thus, pairs of two satellite galaxies in a common halo have a dispersion  $\sqrt{2}$  times larger than pairs of central and satellite galaxies. For two-halo pairs, in which each galaxy comes from a separate halo, the pairwise velocity of the halos contributes to all pairs, but the internal halo dispersions make an additional contribution to pairs involving satellite galaxies. The ratio of satellites to total galaxies is therefore important in determining the behavior of the PVD as a function of luminosity. Although the assumption that central galaxies have no random motions is a working approximation, in tests we find that the results of this assumption are nearly identical to models in which central galaxies have velocities in agreement with recent observational results of van den Bosch et al. (2005), who infer  $\sigma_{\text{cen}} \approx 0.25 \sigma_{\text{sat}}$ . For pairwise statistics these dispersions add in quadrature, thus central motions of this magnitude only increase the central-satellite dispersion by 3%.

To calculate  $w_p(r_p)$  for a given set of HOD parameters, we must first assume a cosmology and a form of the linear matter power spectrum. Because we will use the results of our

<sup>5</sup> All magnitudes quoted in this paper assume  $h = H_0/100 = 1$ . All magnitude bins will be referred to by their lower limit or by a reference number  $i$ , as listed in Table 1.

fitting procedure to populate the halos identified in a collisionless N-body simulation, we take the same values as those used in the simulation. We assume a flat  $\Lambda$ CDM model with  $\Omega_m = 0.3$  and a linear amplitude of fluctuations of  $\sigma_8 = 0.9$ , where  $\sigma_8$  is the variance of matter fluctuations on an  $8 h^{-1}\text{Mpc}$  scale. The shape of the linear matter power spectrum was computed using CMBFAST (Seljak & Zaldarriaga 1996), with  $h = 0.7$  and  $\Omega_b = 0.04$ . The simulation was performed using the Hashed Oct-Tree code of Warren & Salmon (1993). The simulation box is  $400 h^{-1}\text{Mpc}$  per side, with a total of  $1280^3$  particles, giving a mass resolution of  $2.5 \times 10^9 h^{-1} M_\odot$ . Halos are identified in the simulation by the friends-of-friends technique (Davis et al. 1985) with a linking length 0.2 times the mean interparticle separation. For the lowest luminosity,  $M_{b_j} \sim -17$  galaxies, the minimum mass roughly corresponds to a 50-particle halo.

To calculate the model  $w_p(r_p)$  for each bin, the HOD parameters are set to match the abundance of galaxies within each magnitude bin  $\bar{n}_g^i$  as determined by the  $b_j$  luminosity function of Norberg et al. (2002b), with Schechter function parameters  $\Phi_* = 0.0161$ ,  $M_* = -19.7$  and  $\alpha = -1.21$ . This is done by calculating the value of  $M_{\text{min}}^i$  required to match  $\bar{n}_g^i$  once the other parameters for bin  $i$  have been chosen. For high-luminosity bins ( $i \geq 7$ , or  $M_{b_j} \leq -20$ , which we will refer to as the ‘bright’ samples, while referring to  $i < 7$  as the ‘faint’ samples), we allow  $\sigma_{\log M}^i$  to be a free parameter. Recall that the shape of the cutoff in  $\langle N_{\text{cen}} \rangle_M^i$  physically represents the scatter between halo mass and central galaxy luminosity; at bright magnitudes this scatter is expected to be larger due to the steepness of the halo mass function in this regime and the errors in 2dFGRS galaxy magnitudes. For faint samples the scatter is expected to be smaller and magnitude errors do not increase it. We therefore fix  $\sigma_{\log M}^i$  to be 0.15 for these samples. Leaving  $\sigma_{\log M}^i$  as a free parameter for these samples does not result in a statistically better fit to  $w_p(r_p)$ , and in tests we find that the results for the PVD are not sensitive to the value chosen.

We find the best-fitting HOD models using the Markov Chain Monte Carlo (MCMC) technique (see, e.g., Doran & Mueller 2004). In MCMC, an HOD parameter set is randomly chosen from a Gaussian distribution centered on the last accepted parameter set in the MCMC chain. Each element in the chain contains the full set of HOD parameters for all nine magnitude bins. A model is automatically accepted into the chain if it results in a lower  $\chi^2$  value than the previously accepted model, and accepted with the probability  $p = \exp(-\Delta\chi^2/2)$  if the  $\chi^2$  is higher. The number of parameters in our model is 22, consisting of two parameters ( $M_1^i$  and  $M_{\text{cut}}^i$ ) for each  $i = 1 - 6$ , three parameters for each  $i = 7 - 9$  ( $M_1^i$ ,  $M_{\text{cut}}^i$ , and  $\sigma_{\log M}^i$ ), and  $\alpha_{\text{sat}}$ , which is the same for all  $i$ . Because the values of  $M_{\text{min}}^i$  are set by the galaxy number densities, they are not free parameters<sup>6</sup>. The total  $\chi^2$  for each model is the sum of the  $\chi^2$  values obtained for the  $w_p(r_p)$  in each magnitude bin. This large number of parameters makes MCMC an ideal technique for exploring the likelihood of possible solutions and of predictions these solutions make for the luminosity-dependence of the PVD. To check the chain for convergence we use the power spectrum method

<sup>6</sup> We neglect the errors on the galaxy number densities, which are of order a few percent. Due to the steepness of the halo mass function, these errors have little effect on  $M_{\text{min}}^i$ . The dominant source of uncertainty are the  $w_p(r_p)$  measurements.

of Dunkley et al. (2005). The element with the minimum  $\chi^2$  is taken to be our fiducial model, which is listed in Table 1. Although we list the  $\chi^2$  value for each  $w_p(r_p)$  in Table 1, it should be noted that this model was chosen by its total  $\chi^2$ . Therefore, for each individual bin  $i$  there will be models for which  $\chi_i^2$  is lower than that listed in Table 1. We select a subset of 1,000 models from the chain with which to test the constraints on HOD parameters and the PVD inferred from them. Table 1 also lists the galaxy bias factors  $b_g^i$  of the fiducial simulation, as well as the dispersion in  $b_g^i$  from the 1,000 MCMC models. The galaxy bias is straightforward to calculate analytically once the HOD parameters are known, and is determined by

$$b_g^i = \frac{1}{\bar{n}_g^i} \int_0^\infty dM \frac{dn}{dM} \langle N \rangle_M^i b(M), \quad (3)$$

where  $dn/dM$  is the halo mass function (Jenkins et al. 2001) and  $b(M)$  is the halo bias function (Tinker et al. 2005b). The variance of  $b_g^i$  is  $\sim 1 - 2\%$  for most bins, with the tightest constraints for bins  $i = 4 - 7$  ( $-18.5 \geq M_{b_j} \geq -20.0$ ). For lower luminosity bins the limited volume of the samples increases the errors on  $w_p(r_p)$ , while for higher luminosity bins the low number density of galaxies creates significant shot noise in  $w_p(r_p)$ . It should be stressed that the errors in  $b_g^i$  are for a given cosmological model. The galaxy bias obtained from  $w_p(r_p)$  is degenerate with the assumed linear matter power spectrum  $P_{\text{lin}}(k)$ .

Figure 1 shows the occupation functions from the fiducial model for all nine magnitude bins. Panel (a) plots, from left to right,  $\langle N_{\text{sat}} \rangle_M^i$  as a function of  $\log M_h$  for  $i = 1 - 9$ . The thick grey line is the sum of all  $\langle N_{\text{sat}} \rangle_M^i$ , which effectively represents the satellite occupation function for a magnitude threshold sample for galaxies brighter than  $M_{b_j} = -17$ . As expected from both the shape of the galaxy luminosity function and from the increased relative bias as a function of luminosity (see, e.g., Zehavi et al. 2005; Norberg et al. 2002a), the amplitude of the satellite occupation function  $M_1^i$  monotonically increases with the magnitude of the bin. The best-fit values of  $M_{\text{cut}}^i$  are not monotonic with luminosity, and for  $i = 1$  and  $2$  the satellite cutoff occurs at higher mass than  $i = 3$ . Due to the larger errors on  $w_p(r_p)$  for lower luminosity bins, and the fact that  $\langle N_{\text{sat}} \rangle_M^i$  is not well-constrained for occupation numbers significantly less than unity, there is a large range of  $M_{\text{cut}}^i$  allowed by the data for these magnitude bins. We will discuss this further in the following section.

Figure 1b plots  $\langle N_{\text{cen}} \rangle_M^i$  as a function of  $\log M_h$  for the same magnitude bins. The thick grey line, once again showing the sum of all nine  $\langle N_{\text{cen}} \rangle_M^i$  functions, resembles the central occupation function of a magnitude threshold galaxy sample. For the brighter samples, where  $\sigma_{\log M}^i$  is allowed to vary, the shape of the cutoff becomes progressively softer with magnitude. This is expected for two reasons. First, the halo mass function contains an exponential cutoff above  $M_h \sim 10^{13} h^{-1} M_\odot$ , and this steepness is expected to increase the physical scatter between halo mass and central galaxy luminosity. HOD models of the brightest SDSS galaxies by Zehavi et al. (2005) that utilized an exponential cutoff were statistically preferred to models with a step-function form of  $\langle N_{\text{cen}} \rangle_M$ . Second, the magnitude errors in the 2dFGRS are significantly larger than those of SDSS measurements. The exponential cutoff in the galaxy luminosity function creates asymmetric scatter, as galaxies of lower luminosity are scattered into a brighter bin more often

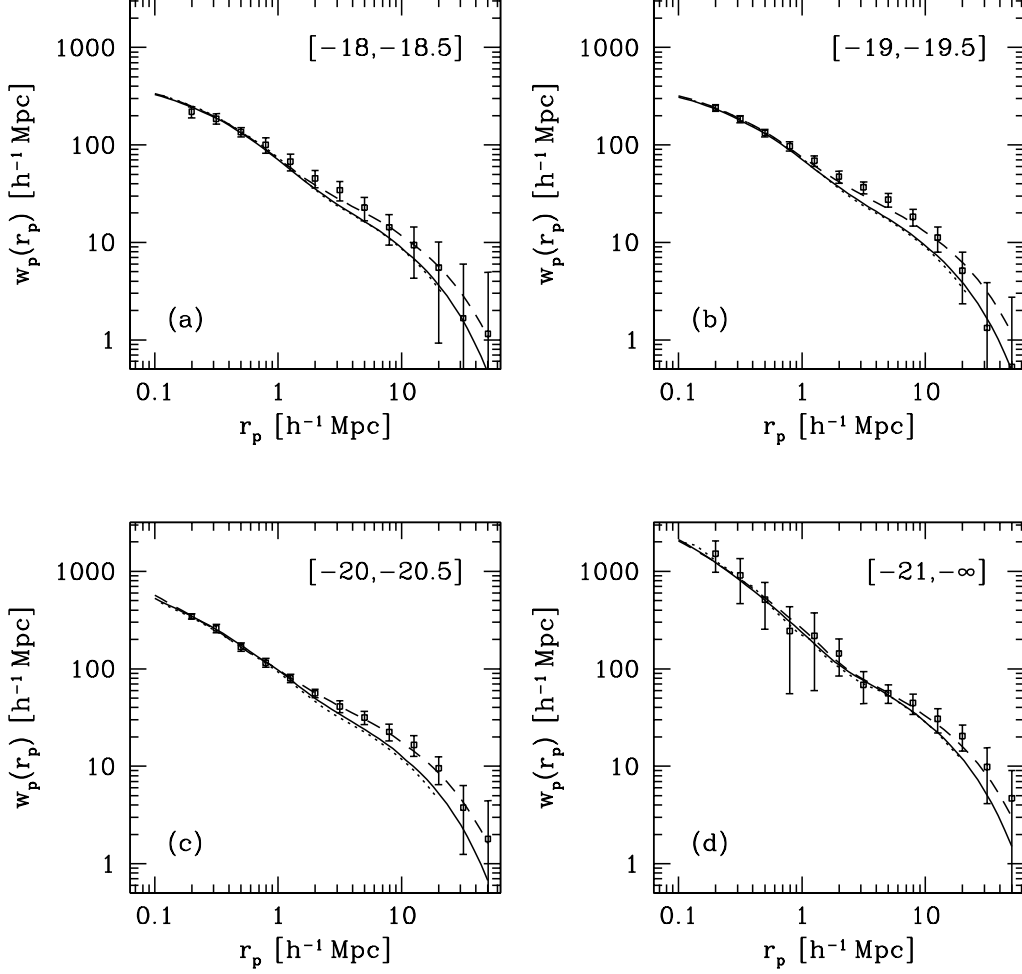


FIG. 2.— Panel (a): Model and observed correlation functions for the galaxies with  $-18 \geq M_{bj} \geq -18.5$  ( $i = 3$  in Table 1). Points with error bars show the measured  $w_p(r_p)$  from the full 2dFGRS data release. The solid line plots the best-fit analytic model, as listed in Table 1. The dotted line shows the  $w_p(r_p)$  obtained from the numerical simulation, once the halos have been populated with the same HOD parameters. The dashed line shows the best-fit HOD when the parameters of Sanchez et al. (2005) are used to calculate the linear matter power spectrum. Panels (b)–(c) are the same as (a), but for magnitude bins  $i = 5, 7$ , and 9.

than higher luminosity galaxies are scattered down.

### 2.3. Mock Galaxy Distributions

To create mock galaxy distributions, we populate the identified halos in the N-body simulation. The number of satellite galaxies for each halo in each magnitude bin is chosen from a Poisson distribution with mean  $\langle N_{\text{sat}} \rangle_M^i$ . The presence of a central galaxy is determined from a nearest integer distribution from the sum of all  $\langle N_{\text{cen}} \rangle_M^i$ , and the luminosity of the central galaxy is chosen from the relative distribution of  $\langle N_{\text{cen}} \rangle_M^i$ . The central galaxies are placed at the center of mass of the halo, while the satellite galaxies are placed randomly throughout the halo according to the universal density profile of Navarro et al. (1997). For each halo mass, the appropriate parameters of the Navarro et al. (1997) profile are calculated for our assumed cosmology from the method of Bullock et al. (2001). The number density profiles of the satellite galaxies are assumed to follow the density profile of the dark matter. As stated above, the central galaxy is given the velocity of the center of mass of the halo, while the satellite galaxies are

given an additional random velocity in each direction drawn from a Gaussian distribution of width

$$\sigma_h^2 = \frac{GM_h}{2R_{200}}, \quad (4)$$

where  $R_{200}$  is the radius at which the mean interior density of the halo is 200 times the universal mean, a value that correlates well on average with the mean densities of friends-of-friends halos with a 0.2 linking length. The relation between  $M_h$  and  $R_{200}$ ,  $R_{200} = (3M_h/4\pi 200\rho_m)^{1/3}$ , makes  $\sigma_h \propto M_h^{1/3}$ . Equation (4) assumes no velocity bias for satellite galaxies. A more general form of equation (4) would have a constant of proportionality  $\alpha_v^2$  relating satellite motions to dark matter (Berlind & Weinberg 2002; Tinker et al. 2005a).

Figure 2 shows the  $w_p(r_p)$  data and the HOD models for four of the magnitude bins, as labeled in each panel. The points with error bars are the 2dFGRS measurements, the solid lines are the fiducial analytic model, and the dotted lines show the measured correlation functions from the populated N-body simulation. The total  $\chi^2$  of this model is 105.3 for

TABLE 1  
HOD PARAMETERS FOR MODELS IN FIGURE 1

$i$	$M_{bj}$	$\chi^2$	$M_{\min}$	$M_{\text{cen}}^{\max}$	$M_1$	$M_{\text{cut}}$	$\sigma_{\log M}$	$b_g$	$\delta_{b_g}$
1	-17.0	15.6	$1.11 \times 10^{11}$	$1.27 \times 10^{11}$	$1.01 \times 10^{13}$	$4.17 \times 10^{12}$	0.15	0.90	0.02
2	-17.5	6.5	$1.57 \times 10^{11}$	$2.01 \times 10^{11}$	$1.08 \times 10^{13}$	$1.07 \times 10^{13}$	0.15	0.92	0.03
3	-18.0	6.3	$2.38 \times 10^{11}$	$2.86 \times 10^{11}$	$1.44 \times 10^{13}$	$5.71 \times 10^{11}$	0.15	0.94	0.015
4	-18.5	10.9	$3.71 \times 10^{11}$	$5.12 \times 10^{11}$	$1.80 \times 10^{13}$	$5.64 \times 10^{12}$	0.15	0.93	0.011
5	-19.0	13.1	$6.73 \times 10^{11}$	$9.15 \times 10^{11}$	$2.61 \times 10^{13}$	$1.84 \times 10^{12}$	0.15	0.96	0.011
6	-19.5	12.5	$1.40 \times 10^{12}$	$2.32 \times 10^{12}$	$3.83 \times 10^{13}$	$1.45 \times 10^{12}$	0.15	1.03	0.007
7	-20.0	9.7	$3.37 \times 10^{12}$	$4.66 \times 10^{12}$	$6.93 \times 10^{13}$	$6.49 \times 10^{12}$	0.09	1.13	0.009
8	-20.5	25.5	$1.22 \times 10^{13}$	$2.67 \times 10^{13}$	$1.82 \times 10^{14}$	$2.19 \times 10^{13}$	0.33	1.29	0.03
9	-21.0	5.2	$6.05 \times 10^{13}$	—	$2.49 \times 10^{14}$	$3.48 \times 10^{14}$	0.56	1.69	0.08

NOTE. — All masses are in units  $h^{-1} M_{\odot}$ . The best-fit value of  $\alpha_{\text{sat}}$  from equation (2) is 1.03.  $M_{\text{cen}}^{\max}$  is the mass at which  $\langle N_{\text{cen}} \rangle_M^i$  is maximum.  $b_g$  is the galaxy bias factor, and  $\delta_{b_g}$  is the dispersion of bias values from the MCMC chain.

$9 \times 12 - 22 = 86$  degrees of freedom, yielding a  $\chi^2$  per degree of freedom of 1.22. But from inspection of Figure 2 it is apparent that the cosmology chosen lacks large-scale power relative to the data. If these data are fit using a  $P_{\text{lin}}(k)$  consistent with the recent analysis of the 2dFGRS large-scale galaxy power spectrum and CMB results (Sanchez et al. 2005), the difference at large scales is ameliorated, and the total  $\chi^2$  is reduced. These fits are shown with the dashed lines in Figure 2. The fits at small scales are unchanged, but at large scales the models match the amplitude of the observations. The  $\chi^2$  per degree of freedom for these fits is 0.61, a value that suggests the possibility that the errors on the data have been overestimated, perhaps by neglecting the covariance in  $w_p(r_p)$  between luminosity bins, or by using too many principal components in the error analysis (Porciani & Norberg 2006). We use the standard  $\Lambda$ CDM fits in order to populate the simulation. Using a simulation with a Sanchez et al. (2005) power spectrum would likely produce small differences in the predicted PVD, but it would not radically change our predictions because the best-fit halo occupation functions are not very sensitive to the assumed power spectrum shape.

#### 2.4. Satellite Fraction of Galaxies

A key quantity for interpreting the results of the HOD analysis is the satellite fraction  $f_{\text{sat}}$ , determined for each bin from the HOD parameters by

$$f_{\text{sat}}^i = \frac{1}{\bar{n}_g^i} \int_0^\infty dM \frac{dn}{dM} \langle N_{\text{sat}} \rangle_M^i. \quad (5)$$

We plot this quantity as well as the different mass scales from the occupation functions in Figure 3. Panel (a) plots the radii of  $M_{\text{cut}}^i$  halos as a function of magnitude for both the fiducial model and the MCMC chain. We present this quantity as a radius rather than a mass because the physical sizes of halos influence the scales at which one-halo and two-halo pairs can contribute. The solid black line represents  $R(M_{\text{cut}}^i)$  for the fiducial model. The shaded region represents the inner 68% of the distribution of  $R(M_{\text{cut}}^i)$  about the median value from the MCMC chain. The distribution of  $R(M_{\text{cut}}^i)$  is calculated for each bin individually, so the upper and lower bounds on the shaded region do not represent a single element in the chain, but the shaded region can be taken as an estimate of the error in  $M_{\text{cut}}^i$ . The bin-to-bin variation seen in the fiducial model is common for individual models from the chain. The midpoint of the shaded area is not defined by a single model, thus the

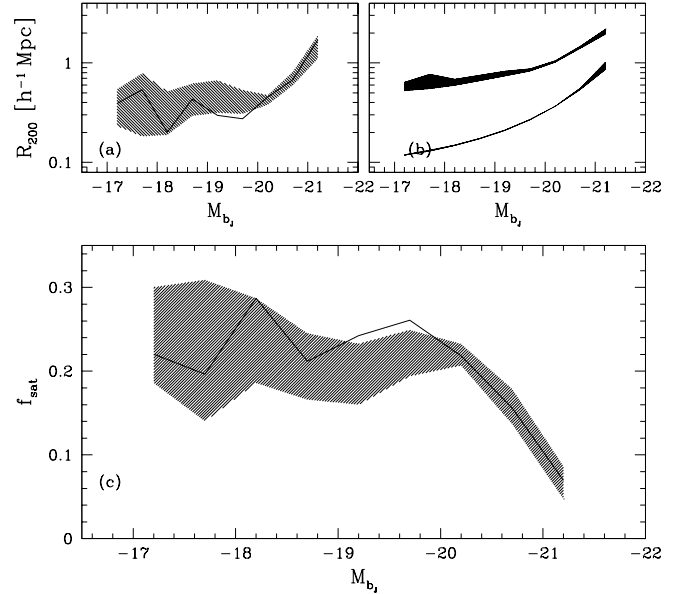


FIG. 3.— Panel (a): Radii of  $M_{\text{cut}}^i$  halos, below which the satellite occupation function is exponentially suppressed, as a function of magnitude. The thick solid line is the relation from the fiducial model, while the shaded region represents the inner 68% of the distribution of  $M_{\text{cut}}^i$  values about the median, taken from the MCMC chain. Panel (b): Radii of  $M_{\text{min}}^i$  halos (bottom shaded region), the cutoff scale for central galaxies, and  $M_{\text{one}}^i$  halos (upper shaded region), which on average contain one satellite. These results are taken from the MCMC analysis only. Panel (c): Fraction of galaxies that are satellites as a function of magnitude. Line and shaded region are as in panel (a).

best-fit model need not follow it. The constraints on  $M_{\text{cut}}^i$  are weakest at lowest luminosities, where the observational errors are largest. The shaded region narrows for  $M_{bj} \leq -19.5$  due to the smaller errors on  $w_p(r_p)$ .

Figure 3b shows the results for  $M_{\text{min}}^i$  and  $M_{\text{one}}^i$  for the MCMC models only, where  $M_{\text{one}}^i$  is the mass at which  $\langle N_{\text{sat}} \rangle_M^i = 1$ , which can be larger than  $M_1^i$  due to the exponential cutoff in equation (2). The shaded regions show the same distribution about the median as in panel (a), with the lower shaded region representing the results for  $M_{\text{min}}^i$  and the upper region representing  $M_{\text{one}}^i$ . The constraints on both these

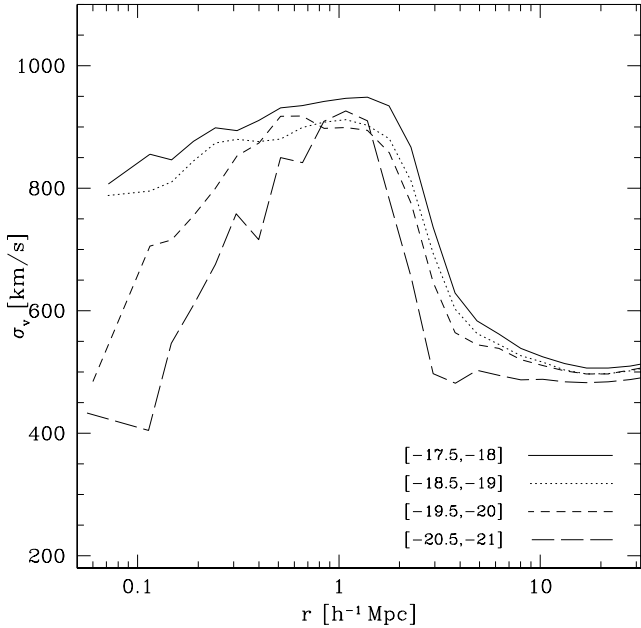


FIG. 4.— Pairwise velocity dispersion as a function of radial separation for four magnitude bins, measured from the populated simulation.

parameters are significantly tighter than those on  $M_{\text{cut}}^i$ . The majority of one-halo pairs come from halos with  $\sim 1$  satellite galaxy, so  $w_p(r_p)$  provides the best constraints on the occupation function around  $M_{\text{one}}^i$ . The value of  $M_{\text{cut}}^i$  influences the shape of  $\langle N_{\text{sat}} \rangle_M^i$  where satellite occupation is typically less than  $\sim 0.5$  and it is therefore less well constrained by  $w_p(r_p)$ . The value of  $M_{\text{min}}^i$  is tightly constrained by both  $w_p(r_p)$  and the galaxy number density. Not plotted is the distribution of  $\alpha_{\text{sat}}$ , which is nearly Gaussian, with a mean of  $\langle \alpha_{\text{sat}} \rangle = 1.03$  and a dispersion of 0.03. It should be noted that this value of  $\alpha_{\text{sat}}$  is degenerate with the assumed value of  $\sigma_8$ ; a lower value of  $\sigma_8$  would increase the inferred value of  $\alpha_{\text{sat}}$  in order to match the observed clustering.

Figure 3c plots  $f_{\text{sat}}^i$  as a function of magnitude for both the fiducial model (solid line) and the MCMC chain (shaded region). As with  $M_{\text{cut}}^i$ , the constraints on  $f_{\text{sat}}^i$  are weaker for fainter samples, where it is constrained to  $\sim 25\%$  about a median value of approximately 0.24. The median value of  $f_{\text{sat}}^i$  drops gradually to  $\sim 0.20$  for  $M_{b_j} \sim M_*$  galaxies, then falls rapidly for samples brighter than  $-20$ , reaching  $f_{\text{sat}}^i = 0.06 \pm 0.02$  for  $i = 9$ . These trends of the satellite fraction are in good agreement with the results of Mandelbaum et al. (2005), who derive satellite fractions from measurements of galaxy-galaxy lensing. Mandelbaum et al. (2005) report  $f_{\text{sat}} \approx 30\%$  for  $L < L_*$  galaxies (in Sloan  $r$ -band) and  $f_{\text{sat}} \approx 0.13$  for galaxies significantly brighter than  $L_*$ .

This estimate of  $f_{\text{sat}}$  depends only weakly on the value of  $\sigma_8$  assumed. To match the observed  $w_p(r_p)$  for a lower value of  $\sigma_8$ ,  $\alpha_{\text{sat}}$  increases to compensate for the lack of high-mass halos. This increases the mean bias of galaxies, but for  $f_{\text{sat}}$  the effects of higher  $\alpha_{\text{sat}}$  and fewer massive halos nearly cancel. In tests where we assume  $\sigma_8 = 0.7$ , the median  $f_{\text{sat}}^i$  is lower by  $\sim 10\%$  than the results in Figure 3, but the 68% confidence regions overlap significantly at all magnitudes. The shape of  $f_{\text{sat}}^i$ , particularly the rapid turnover at  $M_*$ , is also preserved.

### 3. LUMINOSITY AND SCALE DEPENDENCE OF THE PAIRWISE VELOCITY DISPERSION

#### 3.1. Theoretical Predictions

We calculate the PVD from our populated simulation by binning pairs in real space by their radial separation and calculating the line-of-sight pairwise dispersion as a function of  $r$ . We focus on  $\sigma_v$  as a function of three-space separation  $r$  rather than projected separation  $r_p$  for two reasons: presenting  $\sigma_v(r)$  offers a cleaner picture of the PVD at small scales, and the measurements of JB and L05 are intended to be of the PVD at fixed scale and not as function of projected separation. Figure 4 shows the radial dependence of  $\sigma_v$  for four bins of increasing magnitude (samples  $i = 2, 4, 6, 8$  in Table 1). The general shape of the curves is similar to those found for dark matter (e.g., Efstathiou et al. 1988); at large scales,  $\sigma_v$  has little scale dependence, but at  $r \sim 3 h^{-1} \text{Mpc}$ ,  $\sigma_v$  rapidly increases towards smaller  $r$ , peaking at  $r \sim 1 h^{-1} \text{Mpc}$ . This feature represents the transition between two-halo pairs, which are dominated by pairs of central galaxies, and one-halo pairs dominated by pairs of satellite galaxies. At  $r < 1 h^{-1} \text{Mpc}$ , the PVD turns over and monotonically falls with smaller separation. This feature reflects two aspects of the one-halo term. At smaller separation, the relative contribution of central-satellite pairs to the overall pair count increases. Furthermore, as  $r$  decreases, lower-mass halos begin to contribute one-halo pairs, which further reduces the PVD. The former effect is more pronounced for brighter samples because of the smaller fraction of satellite galaxies for these samples.

At a given separation, the trend of  $\sigma_v$  with luminosity can be inferred from Figure 4. For  $r \geq 1 h^{-1} \text{Mpc}$ , the dependence of the PVD on magnitude varies from roughly flat to a moderate decrease with increasing luminosity. At smaller separations, the dependence with luminosity becomes stronger. The location of the transition between one-halo and two-halo pairs also depends on luminosity, with lower-luminosity galaxies transitioning to one-halo pairs at larger separation ( $\sim 3 h^{-1} \text{Mpc}$  for  $i = 2$ ) as compared to bright galaxies ( $\sim 1.5 h^{-1} \text{Mpc}$  for  $i = 8$ ). These results are in good agreement with the analytic HOD calculation of the PVD in Slosar et al. (2006); both models predict the same characteristic shape to  $\sigma_v(r)$ , as well as the same behavior with luminosity at  $r \lesssim 0.5 h^{-1} \text{Mpc}$ . There are minor differences between the predictions, such as the trend of the PVD with luminosity at  $r \gtrsim 1 h^{-1} \text{Mpc}$ , which in their model increases with increasing luminosity, which is the opposite of the trend seen in Figure 4. This discrepancy results from Slosar et al. (2006) fixing the satellite fraction to be a constant value of 0.2 with luminosity. Here we obtain  $f_{\text{sat}}$  directly from the clustering data. Slosar et al. (2006) conclude that the amplitude of the PVD is highly sensitive to the value of  $f_{\text{sat}}$ , and we concur.

We use the results of our fiducial model to interpret the luminosity dependences shown in the previous figure. Figure 5 breaks down the relative contributions of the different types of pairs as a function of magnitude within the radial bins at 0.1, 1, and  $3 h^{-1} \text{Mpc}$ . Within isolated halos, galaxy pairs can be either central-satellite and satellite-satellite. These types of pairs can also exist between two distinct halos, in addition to the central-central contribution. For brevity, we refer to these pair types as cen-sat, sat-sat, and cen-cen in the Figure legend and in the discussion below. Panel (a) plots  $\sigma_v$  for each type of pair at  $r = 0.1 h^{-1} \text{Mpc}$ , as well as the overall PVD with the solid line. At this separation, all galaxy pairs are one-halo. The error bars on the solid line are obtained from

jackknife sampling of the simulation volume into eight octants. Each octant is roughly equal to the volume of the  $M_*$  sample used here. The shaded region once again shows the inner 68% about the median value of  $\sigma_v$  at each luminosity from the MCMC chain. For faint galaxies,  $\sigma_v$  is entirely accounted for by the sat-sat contribution because  $\langle N_{\text{cen}} \rangle_M^i$  and  $\langle N_{\text{sat}} \rangle_M^i$  for these samples do not overlap in mass; halos massive enough to host satellites are too massive to have a low luminosity central galaxy. At magnitudes brighter than  $M_{bj} = -19$ ,  $\sigma_v$  decreases significantly even though the sat-sat PVD remains constant. For the brightest sample the overall PVD is equal to the sat-sat dispersion. Panel (b), which plots the relative contribution of each pair type to the overall pair count as a function of magnitude, shows why this occurs. Below  $M_{bj} = -19$ , sat-sat pairs account for all pairs. At  $\sim M_*$ , the contribution of cen-sat pairs rapidly increases due both to the presence of central galaxies and to the lower fraction of satellites. For  $M_{bj} \leq -21$ , all pairs involve a central galaxy. For this sample,  $\langle N_{\text{sat}} \rangle_M^i = 1$  at  $4 \times 10^{14} h^{-1} M_\odot$ . Although the abundance of halos more massive than this value is low, they do exist in the simulation volume, making it possible to have sat-sat pairs at this magnitude. But because the volume of these halos is relatively large, the probability of having two very bright satellite galaxies within  $0.1 h^{-1} \text{Mpc}$  is low.

Figure 5c plots  $\sigma_v$  for each type of pair at  $r = 1 h^{-1} \text{Mpc}$ . At this separation, which combines both one-halo and two-halo pairs, there are contributions to  $\sigma_v$  from all three pair types. However, the overall PVD lies between the sat-sat and cen-sat dispersions, indicating that one-halo statistics still dominate. This is confirmed in panel (d), which plots the relative fraction of the pair counts. Sat-sat pairs dominate at all magnitudes until the brightest bin. For  $M_{bj} \leq -21$  galaxies, the contribution of cen-sat pairs is equal to sat-sat pairs. Figure 3a showed that the radius of  $M_{\text{cut}}^i$  halos for this sample is  $\gtrsim 1 h^{-1} \text{Mpc}$ , roughly equal to the radial separation shown here. For faint galaxies, cen-sat pairs are more common than pairs between central galaxies, but the abundance of cen-sat pairs gradually decreases with magnitude and cen-cen pairs become more common at  $M_*$ , a result of the lower  $f_{\text{sat}}$  values. Panel (e) plots  $\sigma_v$  for each pair type at  $r = 3 h^{-1} \text{Mpc}$ . At this separation two-halo pairs dominate the statistics, and the overall PVD lies between the cen-sat and cen-cen values. In panel (f), cen-sat and cen-cen pairs are nearly equal for faint galaxies, but there is once again a transition at  $M_*$  where the contribution of cen-cen pairs increases and the other pair types become less common due to the lower satellite fraction.

At small separation, well within the one-halo term, the dependence of the PVD on galaxy luminosity is strong, with brighter galaxies having lower dispersions (with the exception of the brightest galaxies, which live in the highest-mass halos). At intermediate separation, where the transition between one-halo and two-halo pairs occurs, the dependence on luminosity is washed out due to the dominance of satellite pairs at all magnitudes. Outside the one-halo term, but still in the non-linear regime, there is a weak dependence of  $\sigma_v$  on luminosity, falling from  $\sim 700 \text{ km s}^{-1}$  for  $M_{bj} = -17$  to  $\sim 500 \text{ km s}^{-1}$  for  $M_{bj} = -20.5$ . (From Figure 4 it can be seen that this slope becomes shallower at even larger scales.) At all three scales, the results of the MCMC chain demonstrate that these predictions are robust given the errors in the  $w_p(r_p)$  measurements. For the latter two separations, the width of the shaded region is comparable to the jackknife errors from the simulation itself. At face value, these results are at odds with

the observational results of L05 and JB, which show a strong dependence of  $\sigma_v$  on luminosity at  $k = 1 \text{ Mpc}^{-1} h$ , which they argue is equivalent to  $r = 1 h^{-1} \text{Mpc}$ . Our HOD results are also inconsistent with the CLF prediction shown in both papers, which shows a strong trend of increasing PVD with luminosity. In section §4 we will use the JB technique of obtaining the PVD from the redshift-space power spectrum to make a proper comparison of our HOD predictions with these results.

### 3.2. Comparison with 2dFGRS Measurements

Estimates of the PVD from the redshift-space correlation function do not show the sharp transition at  $r \sim 2 h^{-1} \text{Mpc}$  seen in Figure 4. Results from Jing et al. (1998), Zehavi et al. (2002), and Hawkins et al. (2003), from the Las Campanas Redshift Survey, the SDSS, and the 2dFGRS, respectively, show a smooth increase in PVD from large to small scales. However, as discussed in §1, the PVD is not measured directly from the data but inferred by model fitting, and the quantity that is fitted can be interpreted as an “effective” velocity dispersion along the line of sight, calculated as a function of *projected* separation  $r_p$ . In Figure 6, we plot the mean line of sight  $\sigma_v$  as a function of  $r_p$  for pairs with a real-space line of sight separation of less than  $40 h^{-1} \text{Mpc}$ . The four curves are for the same magnitude bins as shown in Figure 4. For comparison, the data from Figure 4 for  $i = 2$  are shown with the open squares. The sharp transition from one-halo to two-halo pairs is smoothed out when plotting  $\sigma_v(r_p)$  because at a given  $r_p$  pairs with larger separations (and lower dispersion) contribute to the average.

Also shown in Figure 6 are the PVD measurements from Hawkins et al. (2003). These data follow the axis on the right-hand side of the plot, the values of which are lower by a factor of 0.8. Thus the amplitude of our HOD prediction is higher than the data. These calculations assume  $\Omega_m = 0.3$  and no velocity bias. Both of these assumptions could be wrong, and there exists growing observational evidence indicating a lower value of  $\Omega_m$  (Bahcall et al. 2003; van den Bosch et al. 2003; Cole et al. 2005; Sanchez et al. 2005; Mohayaee & Tully 2005; Tinker et al. 2005b) and some theoretical evidence for moderate negative velocity bias (Klypin et al. 1999; Berlind et al. 2003; Yoshikawa et al. 2003; Slosar et al. 2006). Therefore we consider the overall normalization of our prediction to be somewhat arbitrary. Read against the right-hand axis, the curves represent models scaled<sup>7</sup> by a factor of  $(0.2/0.3)^{0.55} = 0.8$  to depict a universe with  $\Omega_m = 0.2$ . For fixed power spectrum shape, changing  $\Omega_m$  does not alter the shape of the halo mass function (Zheng et al. 2002). Thus the occupation functions in Figure 1, once scaled to the new value of  $\Omega_m$ , produce identical fits to  $w_p(r_p)$  as in Figure 2. Once scaled to this new value of  $\Omega_m$ , the amplitude and scale dependence of the PVD from the HOD models is generally consistent with the 2dFGRS data. There is some discrepancy at  $r_p = 0.1 h^{-1} \text{Mpc}$ , but this data point suffers from fiber collision effects and incompleteness due to blending of objects in the input catalog. A quantitative comparison between the data and the models is difficult because Hawkins et al. (2003) analyze a flux-limited sample, and we have not replicated the details of their modeling procedure. However, a flux-limited “prediction” would be a linear combination of the different model curves (and the cross-

<sup>7</sup> Halo pairwise velocities scale as  $\Omega_m^{0.6}$  even in the non-linear regime (Zheng et al. 2002), while virial velocities scale as  $\Omega_m^{0.5}$ , thus to a good approximation the pairwise velocities scale as  $\Omega_m^{0.55}$  at all scales.

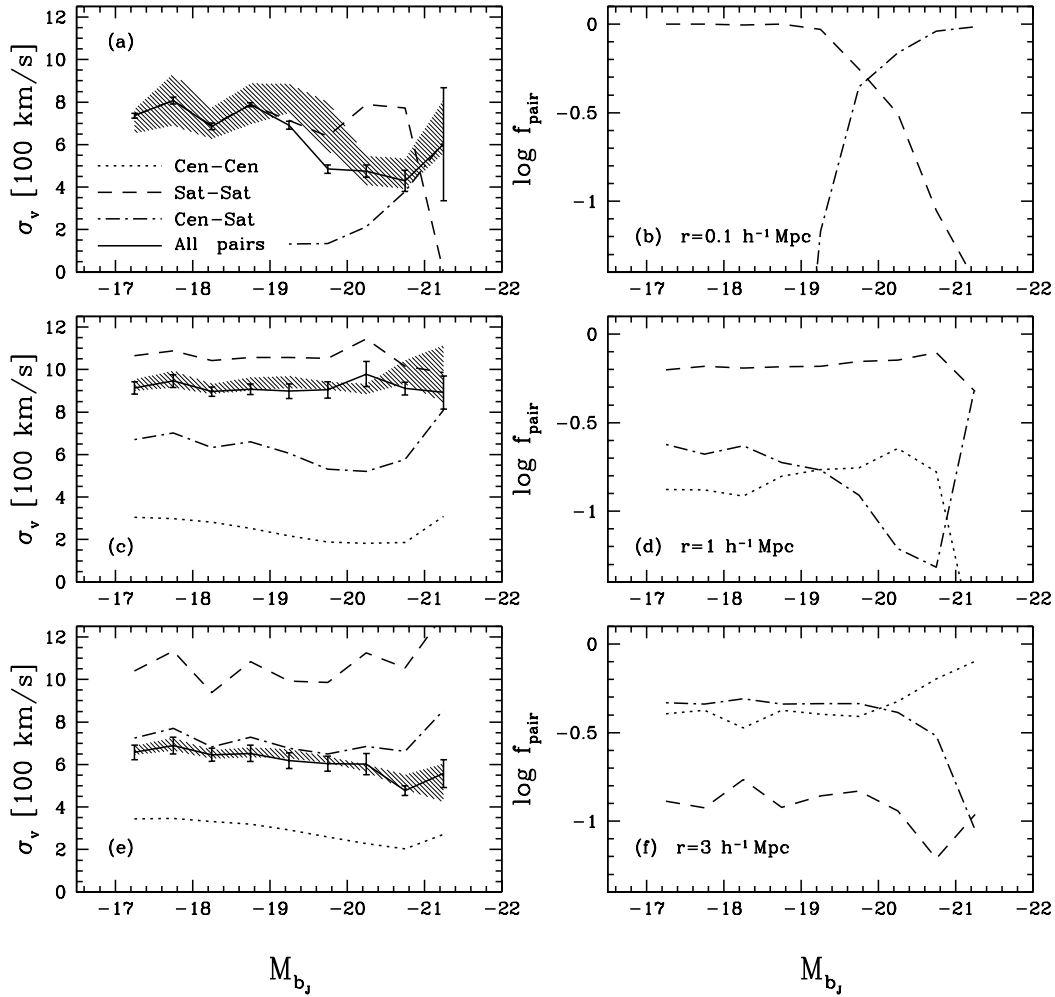


FIG. 5.— Panel (a): PVD at  $r = 0.1 h^{-1} \text{Mpc}$ , broken down by pair type, as a function of luminosity. The dashed line is the PVD for satellite-satellite pairs only, the dash-dotted line represents central-satellite pairs, and the dotted line (not seen in this panel) is for central-central pairs. The solid line is the PVD for all pairs. All curves are obtained from the Fiducial HOD model in Table 1. The shaded region is the inner 68% of the distribution of  $\sigma_v$  values about the median in each magnitude bin, obtained from the MCMC chain. Panel (b): The fraction of the total pair count contributed by each type of pair. Line types are the same as in panel (a). Panels (c) and (d): PVD and pair fraction for each pair type at  $r = 1 h^{-1} \text{Mpc}$ . Line types are as above, with the dotted line representing the central-central contribution. Panel (e) and (f): Same as above, but for  $r = 3 h^{-1} \text{Mpc}$ .

correlated velocity statistics). All the (scaled) model curves show a smooth increase in  $\sigma_v(r_p)$  from  $\sim 400 \text{ km s}^{-1}$  to  $\sim 600 \text{ km s}^{-1}$ , in qualitative agreement with the data.

#### 4. INTERPRETING THE RESULTS OF THE DISPERSION MODEL

JB and L05 measure velocity statistics for volume-limited samples, enabling easier comparison to theoretical predictions. However, the quantity that they measure is not the PVD itself but the parameter  $\sigma_k$  of the dispersion model, which convolves linear theory with an exponential distribution of random galaxy velocities. With this model, the redshift-space power spectrum  $P_Z(k, \mu)$  is related to the real-space power spectrum  $P_R(k, \mu)$  by

$$P_Z(k, \mu) = P_R(k) (1 + \beta \mu^2)^2 (1 + \sigma_k^2 \mu^2 k^2 / 2)^{-1}, \quad (6)$$

where  $\mu$  is the cosine of the angle of the wavevector with the line of sight,  $\beta = \Omega_m^{0.6} / b_g$ , and  $\sigma_k$  is the non-linear galaxy velocity dispersion. The exponential distribution assumed in

equation (6) is both observationally and theoretically motivated (Davis & Peebles 1983; Sheth 1996). In a recent paper, Scoccimarro (2004) presents a thorough critique of the dispersion model, pointing out several deficiencies. First, the physical probability distribution of random velocities implied by equation (6) is not a true exponential dispersion, but in fact contains a discontinuity at the mode. Second, the velocity distribution function of dark matter deviates significantly from an exponential at some scales. This is also true of halo pairwise velocities (Zurek et al. 1994). Inside the scale of the one-halo term, an exponential is a good approximation for galaxies, but in the two-halo term some deviations are to be expected. Third, equation (6) is also only a true convolution if the parameter  $\sigma_k$  is assumed to be constant with scale, but this assumption is contradicted by the results of Figure 4. The scale dependence of the PVD is part what JB and L05 wish to probe, thereby removing one of the theoretical underpinnings of equation (6). In light of this, it is necessary to quantify the accuracy of the JB implementation of the dispersion model in

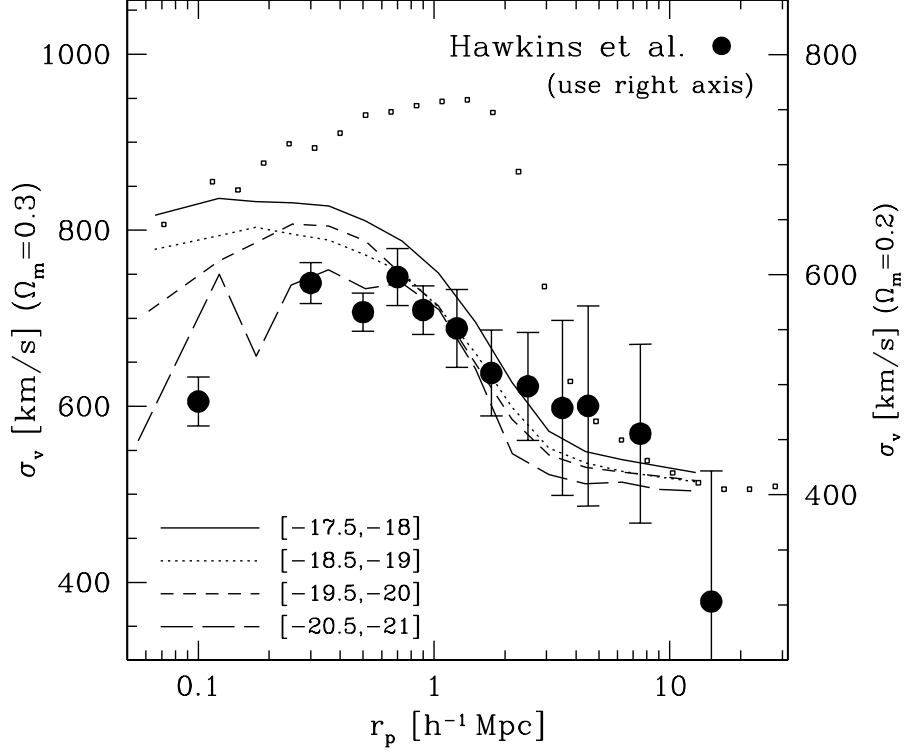


FIG. 6.— Pairwise velocity dispersion as a function of projected separation for the same four magnitude bins shown in the Figure 4 (as a function of 3-d separation). Only pairs with a line-of-sight separation less than  $40 h^{-1}\text{Mpc}$  are counted. The open squares are the results from Figure 4 for the faintest bin shown ( $i=2$ ). The left-hand y-axis is the scale of the PVD as measured from the N-body simulation ( $\Omega_m=0.3$ ). The right-hand y-axis is the scale of the PVD for a universe with  $\Omega_m=0.2$  (see text for details). The filled circles show the 2dFGRS measurements for a flux-limited sample from Hawkins et al. (2003), which should be viewed with respect to the right-hand axis.

order to assess the possible systematic errors accrued through its use. Jing & Börner (2001) test the dispersion model with mock galaxy distributions, finding favorable results, but they apply these tests to multiple realizations of one distribution and not as a function of luminosity. Here we test the dispersion model against all nine magnitude bins.

We calculate the correlation function in redshift space  $\xi(r_p, r_\pi)$ , where  $r_p$  is projected separation and  $r_\pi$  is line-of-sight separation. We measure  $\xi(r_p, r_\pi)$  for  $0 \leq r_p \leq 50 h^{-1}\text{Mpc}$ , and  $-40 \leq r_\pi \leq +40 h^{-1}\text{Mpc}$ . We Fourier transform into  $P_Z(k, \mu)$  by

$$P_Z(k_p, k_\pi) = 2\pi \int dr_\pi \int dr_p \xi(r_p, r_\pi) r_p \cos(k_\pi r_\pi) J_0(k_p r_p) W_g(s) \quad (7)$$

where  $k_p$  and  $k_\pi$  are the projected and line-of-sight directions in  $k$ -space,  $J_0(x)$  is the zeroth-order Bessel function, and

$$W_g(s) = \exp\left(\frac{-s^2}{2S^2}\right), \quad (8)$$

is a weighting function utilized to reduce the influence of large scale distortions, where  $s^2 = r_p^2 + r_\pi^2$ , and  $S = 20 h^{-1}\text{Mpc}$  (see equations [6] and [8] in JB). We calculate the error in  $P_Z(k, \mu)$  from eight jackknife samples and fit equation (6) at each  $k$ , determining the best-fit values of  $P_Z(k, 0)$  and  $\sigma_k$  by  $\chi^2$  minimization. We fix the value of  $\beta$  for each magnitude bin using the value of  $b_g$  from Table 1, but in practice we find the results are insensitive to the value of  $\beta$  (in agreement with JB

and L05).

Figure 7 presents the results of the JB method as applied to our fiducial model. The filled circles plot the dispersion model results as a function of  $k$ , the solid lines plot the true PVD in configuration space (from Figure 4) as a function of  $\pi/r$ , and the dotted lines plot the same curves but as a function of  $1/r$ . The errors on the solid points are the dispersion between the three projections of the simulation volume. For all magnitudes, the dispersion model results show the scale dependence expected from Figure 4; at large scales (small  $k$ ), there is an asymptotic value of  $\sigma_k$ , and at smaller scales the dispersion recovered from the equation (6) rises, reaching a peak between  $k = 1 - 3 \text{ Mpc}^{-1} h$  for most bins. For  $i = 4$  and 5,  $\sigma_k$  has no clear maximum value and continues to increase monotonically with  $k$ . Qualitatively, the dispersion model correctly reproduces the scale-dependence of the pairwise velocity dispersion that is measured in configuration space. In detail, however, the comparison between  $\sigma_k$  and  $\sigma_v$  is problematic. For the lowest three magnitude bins, the transition from two-halo to one-halo pairs is complete by  $k \sim 1 \text{ Mpc}^{-1} h$ , making a favorable comparison with the  $\sigma_v(1/r)$  (dotted) curves, as shown in Jing & Börner (2001). For brighter samples, this transition moves to larger  $k$ , such that for  $i = 6 - 8$ ,  $\sigma_k$  clearly follows the  $\sigma_v(\pi/r)$  (solid) curves. For the other samples, it is not clear that either line best represents the dispersion model results. The results of Figure 7 underscore the difficulty in translating the results of the dispersion model at a given  $k$  to a single physical scale. For faint galaxies, the asymptotic value of  $\sigma_k$  at small  $k$  is larger than the high-separation value of

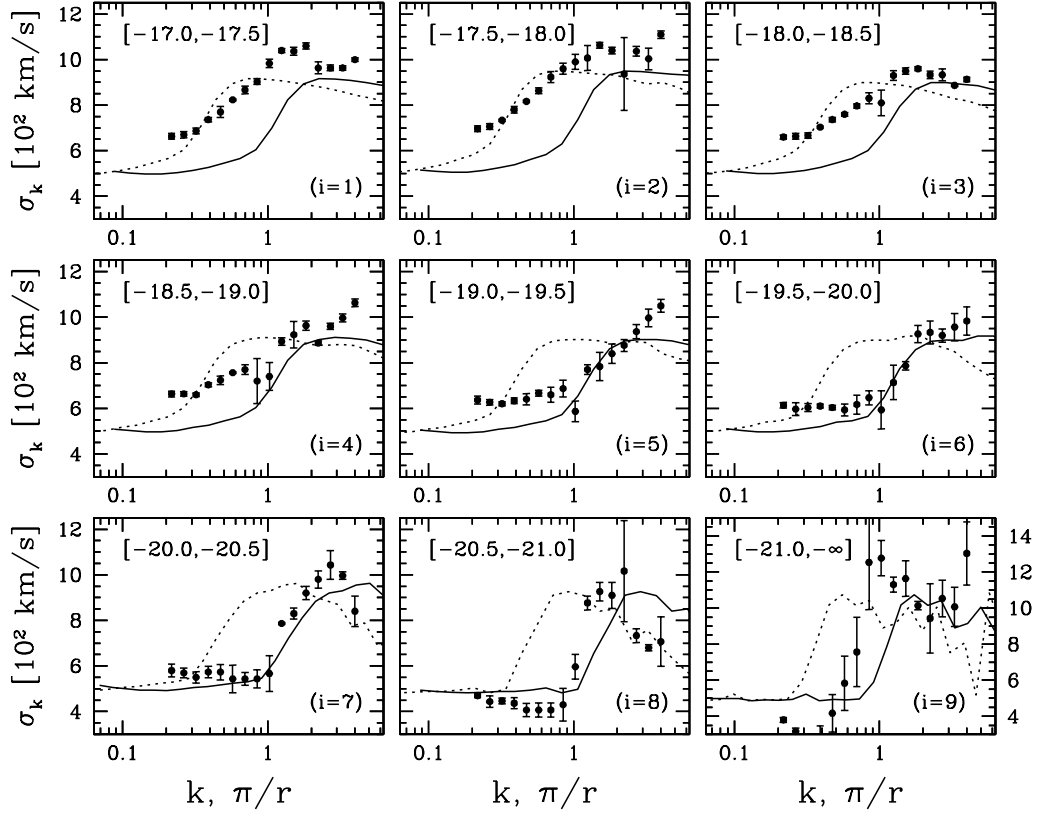


FIG. 7.— Dispersion model results for the fiducial HOD model. Each panel represents a magnitude bin as labeled in top-left corner. Filled circles plot  $\sigma_k$  as a function of  $k$ , and solid lines represent the true PVD measured directly from the simulation (see Figure 4) plotted as a function of  $\pi/r$ . Error bars are the dispersion about the mean of the three projections of the simulation volume. The dotted lines plot the same results as the solid lines, but as a function of  $1/r$ . Note that the range on the y-axis has been increased to a maximum of  $1,500 \text{ km s}^{-1}$  for the bottom right panel ( $i=9$ ), which plots the results for the brightest magnitude bin.

$\sigma_v$  by  $\sim 200 \text{ km s}^{-1}$ , or 40%. For samples near  $M_*$ ,  $\sigma_k$  and  $\sigma_v$  roughly agree at small  $k$ . For the brightest samples,  $\sigma_k$  is smaller than  $\sigma_v$  at large scales.

Figure 8 compares the measurements of JB and L05 at  $k = 1 \text{ Mpc}^{-1} h$  with the dispersion model results of the fiducial model from Figure 7. The open squares and the solid circles are the JB and L05 results, respectively. The results of both JB and L05 show a clear trend of lower velocity dispersion for increasing galaxy luminosity in the range  $M_* + 1 \geq M_{b_j} \geq M_* - 1$ . The claim of JB is that this result is counter-intuitive; brighter galaxies should reside in higher-mass halos and therefore should be subject to larger virial motions. This claim is reinforced by the predictions of a conditional luminosity function model of the 2dFGRS by Yang et al. (2003). The dotted line in Figure 8 plots the predictions of their CLF model, obtained from inspection of Figure 9 in JB and Figure 7 in L05. The CLF prediction is nearly orthogonal to the observational results, with a monotonic trend of increasing  $\sigma_k$  with luminosity.

The large solid squares plot the HOD dispersion model results taken from Figure 7. For several of the magnitude bins, the value of  $\sigma_k$  at  $k = 1$  is slightly off the trend shown by the other points, so we average the model results at  $k = 1$  with the data points on either side. The error bars represent the error in the mean of the nine values (three projections for each of three

values of  $k$ ). In configuration space, there is little to no trend of  $\sigma_v$  with luminosity at  $r = 1 \text{ h}^{-1} \text{ Mpc}$  (Figure 5c). The dispersion model results of the fiducial model show a strong trend with luminosity that is consistent with the dependencies seen in the JB and L05 measurements. Figure 7 demonstrates that at  $k = 1 \text{ Mpc}^{-1} h$ ,  $\sigma_k$  for faint galaxies probes the regime where satellite pairs dominate the statistic, while for brighter galaxies the  $k = 1 \text{ Mpc}^{-1} h$  scale is outside the one-halo regime, where the dispersion is lower. This transition produces an enhanced luminosity dependence in  $\sigma_k$  at this scale. The slope of the HOD results does not appreciably change if we plot  $\sigma_k$  at  $k = 1 \text{ Mpc}^{-1} h$  without averaging over the neighboring  $k$  values.

As in Figure 6, the HOD predictions for this cosmological model are higher than the measurements. The solid line represents the HOD model, scaled again by  $(0.2/0.3)^{0.55} = 0.8$ . This scaled version of our HOD prediction is consistent with the data in both slope and amplitude, even reproducing the sharp upturn of the PVD for the brightest galaxy sample. Jing & Börner (2001) applied the dispersion model to the Las Campanas Redshift Survey, also finding that a model with  $\Omega_m = 0.2$  was a better fit to the data than  $\Omega_m = 0.3$  (for  $\sigma_8 = 1$ ).

The expectation of JB that brighter galaxies should occupy more massive halos is borne out by our HOD modeling. But this creates a strong correlation between the luminosity of

central galaxies and halo mass, while for satellite galaxies the lowest luminosities dominate the distribution (by number) within a given halo. The PVD of faint galaxies is high because of many of them are satellites in high mass halos, and the PVD of luminous galaxies is lower than naively expected because many of them are central galaxies of massive halos that do not inherit the halo velocity dispersion. The results of JB and L05 are thus naturally explainable in the halo occupation context through the effect of  $f_{\text{sat}}$ , discussed in the previous section, and by the systematic effects of using the dispersion model to estimate the PVD.

Why are the predictions of two similar methods (e.g., the HOD and the CLF) so clearly at odds? The real-space clustering results of the JB implementation of the CLF overestimate the observed power spectrum for  $k \gtrsim 1 \text{ Mpc}^{-1} h$  by a factor of 3–5 for galaxies brighter than  $M_{b_j} = -19.5$  (see their Figure 6). This indicates that the number of bright satellite galaxies is also overestimated, leading to more one-halo pairs at these magnitudes than is observed, as well as a higher satellite fraction at these magnitudes. A model that predicts too many bright satellites would predict an increasing PVD with luminosity, as Slosar et al. (2006) demonstrate. In the Yang et al. (2003) CLF, galaxy clustering data are used only at the scale length  $r_0$  of the correlation function, well outside the one-halo term. Therefore, models that overpredict small-scale clustering are not disfavored by high  $\chi^2$  values in their analysis. JB make a similar point about the best-fit parameters of the Yang et al. (2003) model, noting that a model with a steeper faint-end slope of the CLF could be more consistent with the PVD measurements. The differences between our HOD results and the CLF come as some surprise because of the good agreement between the two methods in previous analyses (Tinker et al. 2005b; van den Bosch et al. 2003). These differences are therefore specific to the statistic being modeled here. The utilization of small-scale clustering for determining the occupation functions makes our approach more robust for predicting other statistics at small scales. For statistics that require knowledge of the mean properties of the galaxies within halos, such as the mass-to-light ratio or the overall PVD of galaxies, there should be less bias in CLF results.

## 5. SUMMARY

In this paper we use the halo occupation distribution framework to make robust predictions of the pairwise velocity dispersion of galaxies through modeling of the galaxy two-point correlation function. Our technique is to constrain the parameters of the halo occupation function  $\langle N \rangle_M$  by matching observations of the luminosity-dependent projected correlation function  $w_p(r_p)$ , then populate the halos of an N-body simulation with mock galaxies drawn from the inferred HOD. In our analysis, we make a distinction between central galaxies, which we assume move with the center of mass of the host halo, and satellite galaxies, which we assume have a velocity dispersion within each host halo equal to that of the dark matter. This distinction is crucial to our predictions because the pairs that involve central galaxies have a lower dispersion, so the fraction of satellites strongly influences both the luminosity and scale dependence of the PVD in our predictions. At large separations, the PVD is largely scale independent because halo velocities vary weakly over the scales probed here. At  $r \approx 2 h^{-1} \text{ Mpc}$ , the PVD rapidly increases as satellite-satellite pairs from massive halos dominate. At  $r < 1 h^{-1} \text{ Mpc}$ , the PVD decreases with smaller separation because central-

satellite pairs become more common.

We have focused on the luminosity dependence of the PVD in light of recent claims that the measured luminosity dependence is in conflict with the expectations from halo occupation models. JB and L05 measure a PVD that decreases with luminosity, which is inconsistent with the expectation that brighter galaxies should have higher dispersions because they occupy only high-mass halos. We have shown that this expectation is incorrect; bright galaxies occupy the centers of high-mass halos, so their dispersions are not enhanced by the large virial motions within those halos. The dependence of  $\sigma_v$  on  $M_{b_j}$  is strongly influenced by the fraction of galaxies that are satellites at each magnitude, a quantity that is  $\sim 25\%$  for faint galaxies and rapidly declines for  $M_{b_j} \leq M_*$ . In the CLF prediction for the PVD shown in JB and L05, the PVD monotonically increases with luminosity. This prediction does include the distinction between central and satellite galaxies, but it does not utilize  $w_p(r_p)$  data at small scales to constrain the occupation function, so it cannot accurately constrain  $f_{\text{sat}}$ . While we find that observational uncertainties in  $w_p(r_p)$  and degeneracies among HOD parameters allow some range in our predictions of the PVD, none of our 1,000 MCMC models predict a monotonic increase in  $\sigma_v$  with luminosity.

L05 infer a significant dependence of PVD on color as well as luminosity, with red galaxies having a higher dispersions. L05 conclude that the dependence of the PVD with luminosity results from a population of faint red galaxies in clusters. The large difference in the PVD of faint red and blue galaxies can also be accounted for through halo occupation models. The HOD analysis of SDSS correlation functions in Zehavi et al. (2005) includes analysis of red and blue galaxies. The derived occupation functions predict a higher satellite fraction for red galaxies and a larger number of faint red galaxies in rich systems (see their Figure 23). These occupation functions would produce a substantially higher PVD for red galaxies. The conclusion of L05 is consistent with our interpretation of these data through  $f_{\text{sat}}$ ; faint satellite galaxies are primarily red objects, while faint central galaxies are primarily blue galaxies in low-mass halos.

The PVD is physically interesting quantity, both in terms of cosmology and galaxy formation, but it cannot be measured directly with precision. It can be inferred from the redshift-space correlation function by model fitting, but the results are model dependent and may be affected by the assumption of scale independence of the PVD. We have included a qualitative comparison to the Hawkins et al. (2003) configuration-space estimate of the PVD for a flux-limited sample of 2dFGRS galaxies, but we have focused most of our attention on JB and L05's estimate for volume-limited samples, which use the dispersion model to infer the PVD from the redshift-space power spectrum. We have tested their implementation of this model with our fiducial HOD model, demonstrating that the effective scale in configurations space probed by a given  $k$  scale in Fourier space changes with the luminosity of the sample, complicating the comparison to predicted trends. At low luminosities,  $\sigma_k$  at  $k = 1 \text{ Mpc}^{-1} h$  probes the regime where one-halo pairs dominate the statistics, while for higher luminosities this scale probes the two-halo regime. Therefore, while the luminosity dependence of the PVD in configuration space is flat at  $r = 1 h^{-1} \text{ Mpc}$ ,  $\sigma_k$  at  $k = 1 \text{ Mpc}^{-1} h$  follows a strong trend with luminosity, comparable to the dependence measured by JB and L05.

For our chosen cosmology, with  $\Omega_m = 0.3$  and  $\sigma_8 = 0.9$ , the normalization of the predicted PVD is too high, both in con-

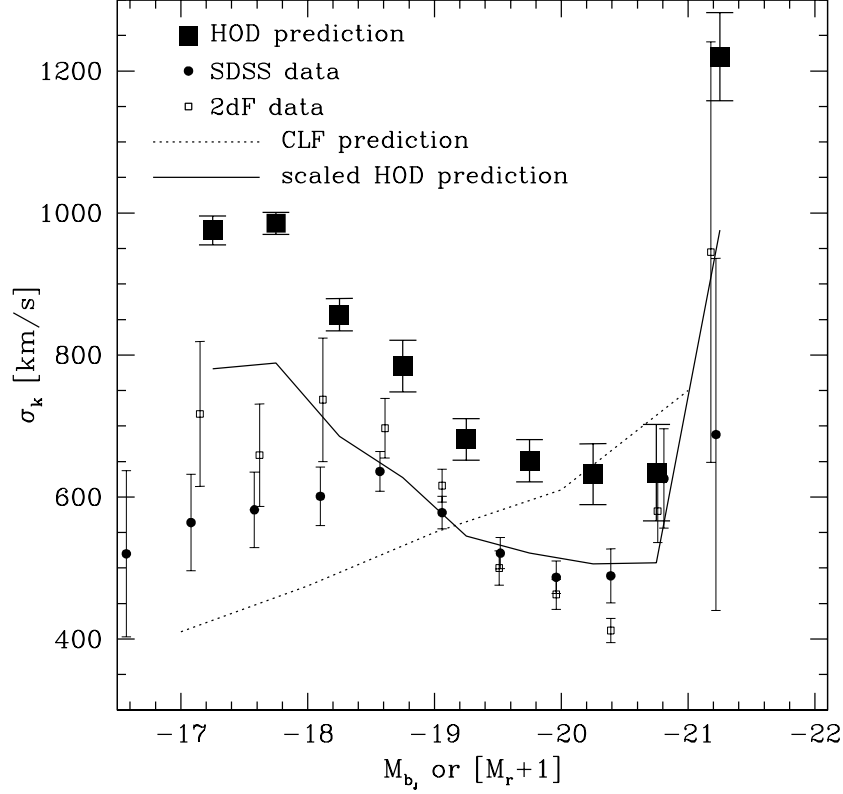


FIG. 8.— Comparison of the fiducial HOD model to the observational results of JB and L05. Open squares and filled circles show the measurements of JB and L05, respectively. The dotted line plots the CLF prediction shown in both papers. The large filled squares are the dispersion model results of the fiducial HOD model, averaged over the three  $k$  values nearest  $k = 1 \text{ Mpc}^{-1} h$ . Error bars are the error in the mean. The solid line shows the HOD results shifted by a factor of 0.8 to represent a universe with  $\Omega_m = 0.2$ .

figuration space and in Fourier space. This discrepancy can be resolved by lowering  $\Omega_m$  to 0.2, thus reducing halo velocities by  $(0.2/0.3)^{0.6}$  and virial dispersions by  $(0.2/0.3)^{0.5}$ . Our predicted PVD would also be lower for a lower value of  $\sigma_8$ . At large scale,  $\sigma_v$  scales roughly linearly with  $\sigma_8$ , while in the one-halo regime  $\sigma_v \sim \sigma_8^{0.5}$  for models constrained to match the same  $w_p(r_p)$ . It is difficult to use the PVD measurements on their own to constrain cosmological parameters, however, because the normalization of our predictions also depends on the velocity bias of both central and satellite galaxies. A more general model would allow the satellite galaxy dispersion to be  $\sigma_{\text{sat}} = \alpha_v \sigma_{dm}$ , and the central galaxy dispersion to be  $\sigma_{\text{cen}} = \alpha_{vc} \sigma_{dm}$ . In principle, both  $\alpha_v$  and  $\alpha_{vc}$  could depend on halo mass and galaxy luminosity because of dynamical effects on the orbits of subhalos that host satellite galaxies. van den Bosch et al. (2005) argue that  $\alpha_{vc} \approx 0.25$  for 2dFGRS groups, but in tests we find that significantly higher values of  $\alpha_{vc}$  are required to alter our predictions. At small scales  $\sigma_v \sim \alpha_v$ , while at large scales  $\sigma_v \sim \alpha_v^{0.5}$ . Thus the normalization of our prediction can be altered by velocity bias. For the  $\Omega_m = 0.3$ ,  $\sigma_8 = 0.9$  model used here,  $\alpha \approx 0.65$  is required to match the large-scale amplitude of the 2dFGRS measurements in Figure 6, a bias that is stronger than that predicted by numerical simulations.

The scale dependence of the PVD predicted by the HOD is in general agreement with that of Hawkins et al. (2003); both the predictions and observations show a smooth transition be-

tween the one-halo and two-halo regimes as a function of projected separation. The  $\sigma_k$  curves in both JB and L05 show an increase in  $\sigma_k$  at  $k \sim 1 \text{ Mpc}^{-1} h$ , but the measurements lack the strong two-halo to one-halo transition in our HOD models (Figure 7). Changes in  $\alpha_v$  and  $\sigma_8$  alter the PVD differently at different scales. The difference in scale dependence may imply a value of  $\alpha_v < 1$ , or a higher  $\sigma_8$  than that assumed here, although a higher value of  $\sigma_8$  would exacerbate the normalization problem in Figures 6 and 8.

Recent papers by Gao et al. (2005) and Harker et al. (2005) have shown that halo formation times correlate with the clustering strength of halos (at fixed mass), implying that formation epoch correlates with local environment. Papers by Wechsler et al. (2005) and Zhu et al. (2006) demonstrate that the number of subhalos correlates with halo formation epoch. The implication for the HOD is that  $\langle N \rangle_M$  could also correlate with environment, an effect not taken into account in our modeling of  $w_p(r_p)$ . If this correlation is strong, the occupation functions derived in Figure 1 will be systematically off their true values. In terms of pairwise velocity statistics, however, these errors would not significantly change the predictions. First, the effect of Gao et al. (2005) and Harker et al. (2005) is strong for low-mass halos, where  $\langle N_{\text{sat}} \rangle_M^i \ll 1$ . Any errors accrued would therefore be localized to the central occupation function. Because the halo-halo velocity dispersion is insensitive to halo mass for masses less than the non-linear mass scale, these types of errors would not alter our predic-

tions. Second, assuming that the satellite occupation functions are systematically off by as much as the difference in  $\langle N_{\text{sat}} \rangle_M^i$  between adjacent magnitude bins, the change in  $\sigma_v$  per half magnitude is  $\sim 5\%$  or less for most separations. Lastly, tests using a hydrodynamic simulation presented in Yoo et al. (2005) demonstrate that any environmental variation of the HOD has little impact on the galaxy correlation function and is not likely to significantly alter the derived occupation functions for luminosity-defined samples. The recent numerical results do underscore the need to quantify the effect that environmental correlations have on predictions of halo occupation models, but we do not expect changes to the qualitative conclusions in this paper.

Ultimately, there is more information about the PVD in the full, two-dimensional redshift-space correlation function than in  $\sigma_v$  from the streaming model or  $\sigma_k$  from the dispersion model. Tinker et al. (2005a) show that modeling redshift-space distortions at small and large scales can separately constrain  $\alpha_v$ ,  $\sigma_8$ , and  $\Omega_m$ . Detailed observational comparisons, imposing the constraint that  $\sigma_8$  and  $\Omega_m$  must be the same for

all samples, could reveal luminosity and mass dependence of  $\alpha_v$ , providing physical insight into the dynamical evolution of galaxies in groups and clusters. We will apply these approaches to 2dFGRS and SDSS observations in future work.

JT would like to thank Shaun Cole, Charlie Conroy, Carlos Frenk, Andrey Kravtsov, Hiranya Peiris, Risa Wechsler, Renbin Yan, Andrew Zentner, and Zheng Zheng for useful discussions. JT would also like to thank the generous hospitality of the Institute for Computational Cosmology at the University of Durham, where part of this work was completed. PN acknowledges support of a PPARC PDRA fellowship and the computer resources of ETH Zurich. DW acknowledges the support of NSF grant AST-0407125. Portions of this work were performed under the auspices of the U.S. Dept. of Energy, and supported by its contract #W-7405-ENG-36 to Los Alamos National Laboratory. Computational resources were provided by the LANL open supercomputing initiative.

## REFERENCES

- Bahcall, N. A., Dong, F., Bode, P., Kim, R., Annis, J., McKay, T. A., Hansen, S., Schroeder, J., Gunn, J., Ostriker, J. P., Postman, M., Nichol, R. C., Miller, C., Goto, T., Brinkmann, J., Knapp, G. R., Lamb, D. O., Schneider, D. P., Vogeley, M. S., & York, D. G. 2003, *ApJ*, 585, 182
- Baugh, C. M., Croton, D. J., Gaztañaga, E., Norberg, P., Colless, M., Baldry, I. K., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Sutherland, W., & Taylor, K. 2004, *MNRAS*, 351, L44
- Bean, A. J., Ellis, R. S., Shanks, T., Efstathiou, G., & Peterson, B. A. 1983, *MNRAS*, 205, 605
- Benson, A. J., Cole, S., Frenk, C. S., Baugh, C. M., & Lacey, C. G. 2000, *MNRAS*, 311, 793
- Berlind, A. A. & Weinberg, D. H. 2002, *ApJ*, 575, 587
- Berlind, A. A., Weinberg, D. H., Benson, A. J., Baugh, C. M., Cole, S., Davé, R., Frenk, C. S., Jenkins, A., Katz, N., & Lacey, C. G. 2003, *ApJ*, 593, 1
- Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001, *MNRAS*, 321, 559
- Cole, S., Fisher, K. B., & Weinberg, D. H. 1994, *MNRAS*, 267, 785
- . 1995, *MNRAS*, 275, 515
- Cole, S., Percival, W. J., Peacock, J. A., Norberg, P., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Cross, N. J. G., Dalton, G., Eke, V. R., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Jenkins, A., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peterson, B. A., Sutherland, W., & Taylor, K. 2005, *MNRAS*, 362, 505
- Colless, M., Dalton, G., Maddox, S., Sutherland, W., Norberg, P., Cole, S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., Couch, W., Cross, N., Deeley, K., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Price, I., Seaborne, M., & Taylor, K. 2001, *MNRAS*, 328, 1039
- Colless, M., Dalton, G., Maddox, S., Sutherland, W., Norberg, P., Cole, S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., Couch, W., Cross, N., Deeley, K., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Price, I., Seaborne, M., & Taylor, K. 2003, *VizieR Online Data Catalog*, 7226, 0
- Cooray, A. & Sheth, R. 2002, *Phys. Rep.*, 372, 1
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, 292, 371
- Davis, M. & Peebles, P. J. E. 1983, *ApJ*, 267, 465
- Doran, M. & Mueller, C. M. 2004, *JCAP*, 0409, 003
- Dunkley, J., Bucher, M., Ferreira, P. G., Moodley, K., & Skordis, C. 2005, *MNRAS*, 356, 925
- Efstathiou, G., Frenk, C. S., White, S. D. M., & Davis, M. 1988, *MNRAS*, 235, 715
- Faltenbacher, A., Kravtsov, A. V., Nagai, D., & Gottlöber, S. 2005, *MNRAS*, 358, 139
- Gao, L., Springel, V., & White, S. D. M. 2005, *MNRAS*, 363, L66
- Hamilton, A. J. S. 1993, *ApJ*, 417, 19
- Harker, G., Cole, S., Helly, J., Frenk, C. S., & Jenkins, A. 2005, *MNRAS*, Submitted (astro-ph/0510488)
- Hatton, S. & Cole, S. 1998, *MNRAS*, 296, 10
- . 1999, *MNRAS*, 310, 1137
- Hawkins, E., Maddox, S., Cole, S., Lahav, O., Madgwick, D. S., Norberg, P., Peacock, J. A., Baldry, I. K., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Jones, B., Lewis, I., Lumsden, S., Percival, W., Peterson, B. A., Sutherland, W., & Taylor, K. 2003, *MNRAS*, 346, 78
- Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., & Yoshida, N. 2001, *MNRAS*, 321, 372
- Jing, Y. P. & Börner, G. 2001, *MNRAS*, 325, 1389
- . 2004, *ApJ*, 617, 782
- Jing, Y. P., Mo, H. J., & Boerner, G. 1998, *ApJ*, 494, 1
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, *MNRAS*, 286, 795
- Klypin, A. A., Gottlöber, S., Kravtsov, A. V., & Khokhlov, A. M. 1999, *ApJ*, 516, 530
- Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., Klypin, A. A., Gottlöber, S., Allgood, B., & Primack, J. R. 2004, *ApJ*, 609, 35
- Landy, S. D. & Szalay, A. S. 1993, *ApJ*, 412, 64
- Li, C., Jing, Y. P., Kauffmann, G., Boerner, G., White, S. D. M., & Cheng, F. Z. 2005, *MNRAS*, accepted, (astro-ph/0509874)
- Ma, C.-P. & Fry, J. N. 2000, *ApJ*, 543, 503
- Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2005, *MNRAS*, submitted, (astro-ph/0511164)
- Mohayaee, R. & Tully, R. B. 2005, *ApJ*, accepted (astro-ph/0509313)
- Navarro, J., Frenk, C., & White, S. 1997, *ApJ*, 490, 493
- Norberg, P., Baugh, C. M., Hawkins, E., Maddox, S., Madgwick, D., Lahav, O., Cole, S., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Lewis, I., Lumsden, S., Peacock, J. A., Peterson, B. A., Sutherland, W., & Taylor, K. 2002a, *MNRAS*, 332, 827
- Norberg, P., Baugh, C. M., Hawkins, E., Maddox, S., Peacock, J. A., Cole, S., Frenk, C. S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Madgwick, D., Peterson, B. A., Sutherland, W., & Taylor, K. 2001, *MNRAS*, 328, 64
- Norberg, P., Cole, S., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Sutherland, W., & Taylor, K. 2002b, *MNRAS*, 336, 907
- Peacock, J. A. & Dodds, S. J. 1994, *MNRAS*, 267, 1020
- Peacock, J. A. & Smith, R. E. 2000, *MNRAS*, 318, 1144
- Peebles, P. J. E. 1976, *ApJ*, 205, L109
- Porciani, C. & Giavalisco, M. 2002, *ApJ*, 565, 24
- Porciani, C. & Norberg, P. 2006, *MNRAS*, submitted
- Sanchez, A. G. et al. 2005, *MNRAS*, in press (astro-ph/0507583)
- Sargent, W. L. W. & Turner, E. L. 1977, *ApJ*, 212, L3
- Scoccimarro, R. 2004, *Phys. Rev. D*, 70, 083007
- Scoccimarro, R., Sheth, R. K., Hui, L., & Jain, B. 2001, *ApJ*, 546, 20
- Seljak, U. 2000, *MNRAS*, 318, 203
- Seljak, U. & Warren, M. S. 2004, *MNRAS*, 355, 129
- Seljak, U. & Zaldarriaga, M. 1996, *ApJ*, 469, 437
- Sheth, R. K. 1996, *MNRAS*, 279, 1310
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, 323, 1
- Sheth, R. K. & Tormen, G. 1999, *MNRAS*, 308, 119

- Slosar, A., Seljak, U., & Tasitsiomi, A. 2006, MNRAS, 366, 1455
- Tinker, J. L., Weinberg, D. H., & Zheng, Z. 2005a, MNRAS, submitted (astro-ph/0501029)
- Tinker, J. L., Weinberg, D. H., Zheng, Z., & Zehavi, I. 2005b, ApJ, 631, 41
- van den Bosch, F. C., Mo, H. J., & Yang, X. 2003, MNRAS, 345, 923
- van den Bosch, F. C. et al. 2005, MNRAS, submitted (astro-ph/0502466)
- Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2005, ApJL, submitted (astro-ph/0506395)
- Warren, M. S. & Salmon, J. K. 1993, in Supercomputing '93
- Wechsler, R. H., Zentner, A. R., Bullock, J. S., & Kravtsov, A. V. 2005, ApJ, submitted, (astro-ph/0512416)
- Yang, X., Mo, H. J., & van den Bosch, F. C. 2003, MNRAS, 339, 1057
- Yoo, J., Tinker, J. L., Weinberg, D. H., Zheng, Z., Katz, N., & Davé, R. 2005, ApJ, submitted, (astro-ph/0511580)
- Yoshikawa, K., Jing, Y. P., & Börner, G. 2003, ApJ, 590, 654
- Zehavi, I., Blanton, M. R., Frieman, J. A., Weinberg, D. H., Mo, H. J., Strauss, M. A., et al. 2002, ApJ, 571, 172
- Zehavi, I., Weinberg, D. H., Zheng, Z., Berlind, A. A., Frieman, J. A., Scoccimarro, R., Sheth, R. K., Blanton, M. R., Tegmark, M., Mo, H. J., et al. 2004, ApJ, 608, 16
- Zehavi, I., Zheng, Z., Weinberg, D. H., Frieman, J. A., Berlind, A. A., Blanton, M. R., Scoccimarro, R., Sheth, R. K., Strauss, M. A., Kayo, I., Suto, Y., Fukugita, M., Nakamura, O., Bahcall, N. A., Brinkmann, J., Gunn, J. E., Hennessy, G. S., Ivezić, Ž., Knapp, G. R., Loveday, J., Meiksin, A., Schlegel, D. J., Schneider, D. P., Szapudi, I., Tegmark, M., Vogeley, M. S., & York, D. G. 2005, ApJ, 630, 1
- Zheng, Z. 2004, ApJ, 610, 61
- Zheng, Z., Berlind, A. A., Weinberg, D. H., Benson, A. J., Baugh, C. M., Cole, S., Davé, R., Frenk, C. S., Katz, N., & Lacey, C. G. 2005, ApJ, 633, 791
- Zheng, Z., Tinker, J. L., Weinberg, D. H., & Berlind, A. A. 2002, ApJ, 575, 617
- Zhu, G., Zheng, Z., Lin, W. P., Jing, Y. P., Kang, X., & Gao, L. 2006, ApJ, submitted, (astro-ph/0601120)
- Zurek, W. H., Quinn, P. J., Salmon, J. K., & Warren, M. S. 1994, ApJ, 431, 559